



SUBSTITUTE SEQUENCE LISTING

GENERAL INFORMATION:

(i) APPLICANT: JONATHAN EDWARD LIGHTNER
JOHN JOSEPH OKULEY

(ii) TITLE OF INVENTION: GENES FOR MICROSOMAL
FATTY ACID DELTA-12
DESATURASES AND RELATED
ENZYMES FROM PLANTS

(iii) NUMBER OF SEQUENCES: 61

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: E. I. DU PONT DE NEMOURS
AND COMPANY
(B) STREET: 1007 MARKET STREET
(C) CITY: WILMINGTON
(D) STATE: DELAWARE
(E) COUNTRY: U.S.A.
(F) ZIP: 19898

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: Macintosh
(C) OPERATING SYSTEM: Macintosh 6.0
(D) SOFTWARE: Microsoft Word 4.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) OFFICE APPLICATION DATA:

(A) APPLICATION NUMBER: U.S. 07/977,339
(B) FILING DATE: 11-NOV-1992

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Siegell, Barbara C.
(B) REGISTRATION NUMBER: 36,684
(C) REFERENCE/DOCKET NUMBER: BB-1043-B

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (302) 992-4227
(B) TELEFAX: (302) 992-7949
(C) TELEX: 31420

SEQUENCE INFORMATION:

(i) POLYNUCLEOTIDE SEQUENCE LISTING:
(A) LENGTH: 1372 base pairs
(B) TYPE: genomic DNA
(C) STRANDELNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

(B) CLONE: p32103

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 33..1244

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGAGAGAGAG ATTCTGGGAA GGAGCTTCTT CTTCTAGGG TGTTCATCGT TATTAACGTT	60
ATCGCCCTA CGTCAGCTCC ATCTCCAGAA AC ATG GGT GCA GGT GGA AGA ATG Met Gly Ala Gly Gly Arg Met	113
1 5	
CCG GTT CCT ACC TCT TCC AAG AAA TCG GAA ACC GAC ACC ACA AAG CGT Pro Val Pro Thr Ser Ser Lys Lys Ser Glu Thr Asp Thr Thr Lys Arg	161
10 15 20	
GTC CCG TGC GAG AAA CCG CCT TTC TCG GTG GGA GAT CTG AAG AAA GCA Val Pro Cys Glu Lys Pro Pro Phe Ser Val Glu Asp Leu Lys Lys Ala	209
25 30 35	
Ala Ile Val Pro His Tyr Phe Lys Arg Ser Ile Pro Arg Ser Phe Ser Tyr 40 45 50 55	255
CTT ATC AGT GAC ATC ATT ATA GCC TCA TGC TTC TAC TAC GTC GCC ACC Leu Ile Ser Asp Ile Ile Ala Ser Cys Phe Tyr Tyr Val Ala Thr	305
60 65 70	
AAT TAC TTC TCT CTC CCT CAG CCT CTC TCT TAC TTG GCT TGG CCA Asn Tyr Phe Ser Leu Leu Pro Gln Pro Leu Ser Tyr Leu Ala Trp Pro	355
75 80 85	
GTC TAT TGG GCG TGT CAA GGC TGT GTC CTA ACT GGT ATC TGG GTC ATA Leu Tyr Trp Ala Cys Gln Gly Cys Val Leu Tyr Gly Ile Trp Val Ile	401
90 95 100	
GGC CAC GAA TGC GGT CAC CAA GCA TTT AGC GAC TAC CAA TGG CTG GAT Ala Glu Glu Tyr Glu Glu Asp Pro Ile Asp Tyr Glu Ile Asp	449
105 110 115	
GAC AGA GAT GGT CCT ATC TTT GAT TCA TGC CTG CTC GCT TAT TTT Asn Arg Glu Gln Tyr Cys Ile Asp Glu Phe Glu Cys Glu Ile Tyr ...	467
120 125 130 135	
TGC TGG AAG TAT AGT CAT TGT CCT CAC CAT TCC AAG ACT GGA TGC CCT Ser Trp Lys Tyr Ser His Arg Arg His Gln Ser Asn Thr Glu Ser Leu	545
140 145 150	
GAA AGA GAT GAA GTA TTT GTC GCA AAG CAG AAA TCA GCA ATC AAG TGG Gln Arg Asp Glu Val Phe Val Phe Lys Gln Lys Ser Ala Ile Lys Tyr	593
155 160 165	

TAC	GCC	AAA	TAC	CTC	AAC	AAC	CCT	CTT	GGA	CGC	ATC	ATG	ATG	TTA	ACC		641	
Tyr	Gly	Lys	Tyr	Leu	Asn	Asn	Pro	Leu	Gly	Arg	Ile	Met	Met	Leu	Thr			
170				175							180							
GTC	CAG	TTT	GTC	CTC	GCC	TGG	CCT	TTG	TAC	TTA	GCC	TTT	AAC	GTC	TCT		689	
Ile	Gln	Phe	Val	Leu	Gly	Trp	Pro	Leu	Tyr	Leu	Ala	Phe	Asn	Val	Ser			
185				190							195							
GCC	AGA	CGG	TAT	GAC	GCC	TTC	GCT	TGC	CAT	TTT	TTC	CCC	AAC	GCT	CCC		737	
Gly	Arg	Pro	Tyr	Asp	Gly	Phe	Ala	Cys	His	Phe	Phe	Pro	Asn	Ala	Pro			
200				205						210				215				
ATC	TAC	AAT	GAC	GCA	GAA	CGC	CTC	CAG	ATA	TAC	CTC	TCT	GAT	GCG	GGT		745	
Ile	Tyr	Asn	Asp	Arg	Glu	Arg	Leu	Gln	Ile	Tyr	Leu	Ser	Asp	Ala	Gly			
220				225						230				235				
ATP	CTA	GCC	GTC	TGT	TTT	GGT	CTT	TAC	CGT	TAC	GCT	GCT	GCA	CAA	GGS		533	
Ile	Leu	Ala	Val	Cys	Phe	Gly	Leu	Tyr	Arg	Tyr	Ala	Ala	Ala	Gln	Gly			
235				240						245								
ATG	GCC	TGG	ATG	ATC	TGC	CTC	TAC	GGA	GTA	CGG	CTT	CTG	ATA	GTG	AAT		581	
Met	Ala	Ser	Met	Ile	Cys	Leu	Tyr	Gly	Val	Pro	Leu	Leu	Ile	Val	Asn			
250				255						260								
GCG	TTC	CTG	GTC	TTG	ATC	ACT	TAC	TTG	CAG	CAC	ACT	CAT	CCC	TGG	TTG		539	
Ala	Phe	Leu	Val	Leu	Ile	Tyr	Leu	Gln	His	Thr	Bis	Pro	Ser	Leu				
265				270						275								
CCT	CAC	TAT	GAT	TCA	TCA	GAG	TGG	GAC	TGG	CIC	AGG	GGA	GCI	TTG	GCT		577	
Pro	His	Tyr	Asp	Ser	Ser	Gly	Trp	Asp	Trp	Leu	Arg	Gly	Ala	Leu	Ala			
280				285						290				295				
ACC	GTA	GAT	AGA	GAC	TAC	GGA	ATC	TTG	AAC	AAG	GTG	TTC	CAC	AAC	ATT		1025	
Thr	Val	Asp	Arg	Asp	Tyr	Gly	Ile	Leu	Asn	Lys	Val	Phe	His	Asn	Ile			
300				305						310								
ACA	GAC	ACA	CAC	GTG	GCT	CAT	CAC	CTG	TTC	TGG	ACA	ATG	CCG	CAT	TAT		1073	
Thr	Asp	Thr	His	Val	Ala	His	His	Leu	Phe	Ser	Thr	Met	Pro	His	Tyr			
315				320						325								
AAC	GCA	ATG	GAA	GCT	ACA	AAG	GCG	ATA	AAG	CCA	ATT	CTG	GGA	GAC	TAT		1121	
Asn	Ala	Met	Glu	Ala	Thr	Lys	Arg	Ile	Lys	Pro	Ile	Leu	Gly	Asp	Tyr			
330				335						340								
"I"	"A"	"T"	"C"	"G"	"G"	"A"	"G"	"T"	"A"	"G"	"T"	"A"	"G"	"G"	"A"		1048	
"Tyr"	"Phe"	"Asp"	"Ser"	"Thr"	"Trp"	"Ile"	"Tyr"	"Val"	"Ala"	"Met"	"Tyr"	"Arg"	"Gln"	"Ala"	"Ile"			
345				350						355				360				
"G"	"A"	"G"	"T"	"C"	"T"	"G"	"A"	"G"	"T"	"A"	"G"	"G"	"A"	"G"	"T"		1049	
"Lys"	"Glu"	"Cys"	"Ile"	"Tyr"	"Val"	"Glu"	"Pro"	"Asp"	"Arg"	"Glu"	"Gly"	"Asp"	"Lys"	"Lys"	"Gly"			
365				370						375				380				
"Tyr"	"Trp"	"Val"	"Trp"	"Asp"	"AAT"	"AAC"	"TTA"	"TGACCATGAT"	"GCTGAAAGAAA"	"TTGTGGACCT"							1271	
"Val"	"Tyr"	"Trp"	"Tyr"	"Asn"	"Asn"	"Lys"	"Leu"											
380																		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		1272
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"	"TTT"		
"TTT"	"TTT"	"TTT"	"TTT"	"TTT"</														

CCATTTGTT GTGTTATGAC ATTTTGGCTG CTCATTATGT T

1372

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 383 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser
1 5 10 15

Glu Thr Asp Thr Thr Lys Arg Val Pro Cys Glu Lys Pro Pro Phe Ser
20 25 30

Val Gly Asp Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
35 40 45

Ile Pro Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile Ile Ile Ala Ser
50 55 60

Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
65 70 75 80

Leu Ile Tyr Ala Ile Ile Ser Lys Ile Ala Cys Glu Gly Cys Val
85 90 95

Leu Thr Gly Ile Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
100 105 110

Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser
115 120 125

Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
130 135 140

His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
145 150 155 160

Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu
165 170 175

Gly Arg Ile Met Met Ile Thr Val Ser Ile Val Leu Arg Ile Pro Leu
180 185 190 195

Asn Lys Ser Ala Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu
195 200 205

His Phe Phe Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln
210 215 220 225

Ile Tyr Leu Ser Asp Ala Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr
225 230 235 240

Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser Met Ile Cys Leu Tyr Gly

345	250	255	
Val Pro Leu Leu Ile Val Asn Ala Phe Leu Val Leu Ile Thr Tyr Leu 260	265	270	
Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp 275	28	285	
Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu 290	295	300	
Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu 305	310	315	320
Phe Ser Thr Met Pro His Tyr Asn Ala Met Glu Ala Thr Lys Ala Ile 325	330	335	
Lys Pro Ile Leu Gly Asp Tyr Tyr Gin Phe Asp Gly Thr Pro Trp Tyr 340	345	350	
Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp 355	360	365	
Arg Glu Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn Asn Lys Leu 370	375	380	

(2) INFORMATION FOR SEQ ID NO:3:

(i) MOLECULE CHARACTERISTICS:

- (A) LENGTH: 1426 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ORIGINAL SOURCE:

(A) ORGANISM: Brassica napus

(v) FEATURE:

- (A) NAME KEY: CDS
- (B) LOCATION: 130..1264

(vi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

GGCACCGAGCT GATTCGCAAT TGCACACAG AGGAGACACA GAGAGAGTTT GAAAGGGAGC' 60
TTCCTTGTA GGTGATCGT TATTAACGTT AAATCTTCAT CCCCCCTAC GTCAGGCCAGC 120
TCAAGAAAC ATG GGT GCA GGT GGA AGA ATG CAA GTG TCT CCT CCC TCC      168
    Met Gly Ala Gly Gly Arg Met Glu Val Ser Pro Ile Ser
    1           5           10
AAA AAG TCT GAA ACC GAC AAC ATC AAG CGC GTA CCC TGC GAG ACA CCG      216
Lys Lys Ser Glu Thr Asp Asn Ile Lys Arg Val Pro Cys Glu Thr Pro
    15          20          25
'CC TT' ATT ST' AGA GAA TTC AAG AAA GCA ATC CCA CGG CAC TGT TTC      264

```

Phe	Phe	Thr	Val	Gly	Glu	Leu	Lys	Ala	Ile	Pro	Pro	His	Cys	Phe	
34					35				40					45	
AAA	CGC	TCG	ATC	CCT	CGC	TCT	TTC	TCC	TAC	CTC	ATC	TGG	GAC	ATC	ATC
Lys	Arg	Ser	Ile	Pro	Arg	Ser	Phe	Ser	Tyr	Leu	Ile	Trp	Asp	Ile	Ile
					50				55				60		
ATA	GCC	TCC	TGC	TTT	TAC	TAC	GTC	GCC	ACC	ACT	TAC	TTC	GCT	CTG	CTC
Ile	Ala	Ser	Cys	Phe	Tyr	Tyr	Val	Ala	Thr	Thr	Tyr	Phe	Pro	Leu	Leu
					65				70				75		
CCT	SAC	CCT	CTC	TCC	TAC	TTC	GCC	TGG	CCT	CTC	TAC	TGG	GCC	TGC	CAG
Pro	His	Pro	Leu	Ser	Tyr	Phe	Ala	Trp	Pro	Leu	Tyr	Trp	Ala	Cys	Gln
					80				85				90		
GGC	TGC	GTC	CTA	ACC	GGC	GTC	TGG	GTC	ATA	GCC	CAC	GAG	TGC	GGC	CAC
Gly	Cys	Val	Leu	Thr	Gly	Val	Trp	Val	Ile	Ala	His	Glu	Cys	Gly	His
					95				100				105		
CAC	GCC	TTC	AGC	GAT	TAC	CAC	TGG	CTG	GAT	GAC	ACC	GTC	GGC	CTG	ATP
His	Ala	Phe	Ser	Asp	Tyr	Gln	Trp	Leu	Asp	Asp	Thr	Val	Gly	Leu	Ile
					110				115				120		125
TTC	CAC	TCC	CTC	CTG	GTC	CCT	TAC	TTC	TCC	TGG	AAG	TAC	AGT	CAT	
Phe	His	Ser	Phe	Leu	Leu	Val	Pro	Tyr	Phe	Ser	Trp	Lys	Tyr	Ser	His
					130				135				140		
CCA	TGC	CAT	CAT	TCT	ATC	ATC	GGC	TGC	GTC	GTC	AGA	GAA	GAA	GTC	TTT
Ala	Cys	Asp	Asp	Leu	Leu	Leu	Gly	Ser	Leu	Leu	Gly	Gly	Gly	Gly	Leu
					145				150				155		
GTC	CCT	AAG	AAG	AAG	TCA	GAC	ATC	AAG	TGG	TAC	GGC	AAG	TAC	CTG	AAC
Val	Pro	Lys	Lys	Lys	Ser	Asp	Ile	Lys	Trp	Tyr	Gly	Lys	Tyr	Leu	Asn
					160				165				170		
AAU	CCT	TTC	GGA	CGC	ATC	GTC	ATG	TTC	AGC	GTT	CAG	TTC	ACT	CTT	GGG
Asn	Pro	Leu	Gly	Arg	Ile	Val	Met	Ile	Thr	Val	Gln	Phe	Thr	Leu	Gly
					175				180				185		
TGG	CCT	TTC	TAC	TTA	GCC	TTC	AAC	GTC	TGG	GGG	AGA	CCT	TAC	GAT	GGG
Trp	Pro	Leu	Tyr	Ile	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	Asp	Gly
					190				195				200		205
GGG	TTC	GCT	TGC	CAT	TTC	CAC	CCC	AAC	CCT	TTC	ATC	TAC	AAU	GAT	CGT
Gly	Phe	Ala	Cys	His	Phe	His	Pro	Asn	Ala	Pro	Ile	Tyr	Asn	Asp	Arg
					210				215				220		
GAG	GGT	CTT	CAG	ATA	TAT	ATC	TTC	CAC	GTC	TCT	ATC	TTC	TTT	ATC	TTG
Asp	Gly	Leu	Val	Leu	Tyr	Ile	Leu								
					225				230				235		
TAT	TGT	TAC	GCG	TAC	GCT	GCT	CAA	GAA	GTC	GCC	TGG	ATG	GTC		
Tyr	Gly	Leu	Tyr	Arg	Tyr	Ala	Ala	Val	Gly	Val	Ala	Ser	Met	Val	
					240				245				250		
TGC	TTC	TAC	GGA	GTT	CTT	CTG	ATT	GTC	AAC	GGG	TTC	TTA	GTT	TTG	
Cys	Phe	Tyr	Gly	Val	Pro	Leu	Ile	Val	Asn	Gly	Phe	Leu	Val	Leu	
					255				260				265		

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 384 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Cys Phe Tyr Tyr Val Ala Thr Thr Tyr Phe Pro Leu Leu Pro His Pro
 65 70 75 80

Leu Ser Tyr Phe Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val
 85 90 95

Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
 100 105 110

Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser
 115 120 125

Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
 130 135 140

His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
 145 150 155 160

Lys Lys Ser Asp Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu
 165 170 175

Gly Arg Thr Val Met Leu Thr Val Gln Phe Thr Leu Gly Trp Pro Leu
 180 185 190

Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Gly Phe Ala
 195 200 205

Cys His Phe His Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu
 210 215 220

Gln Ile Tyr Ile Ser Asp Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu
 225 230 235 240

Tyr Arg Tyr Ala Ala Val Gln Gly Val Ala Ser Met Val Cys Phe Tyr
 245 250 255

Gly Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val Leu Ile Thr Tyr
 260 265 270

Leu Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp
 275 280 285

Asp Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile
 290 295 300

Leu Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His
 305 310 315

Leu Pro Ser Thr Met Pro His Ile Pro Ala Met Glu Ile Thr Ile Ala
 320 325 330 335

Pro Lys Pro Ile Leu Gly Glu Ily Tyr Gln Phe Asp Gly Thr Pro Val
 340 345 350

Val Lys Ala Met Trp Arg Glu Ala Lys Glu Cys Ile Tyr Val Gln Pro
 355 360 365

Asp Arg Gln Gly Glu Lys Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu
 370 375 380

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS
(A) LENGTH: 146.2 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: ssDNA

HYPOTHETICAL; NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL NOTICE:

(A) ORGANISM: Glycine max

(vii) IMMEDIATE SOURCE:

(B) CLONE: pSF1-165K

(ix) FEATURE:

A) NAME / KEY: ODS

B) LOCATION: 105.1247

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCATATACTA AATTGTTTGT ATTGGTGTTG CCCCTTCGGT CCCAAAGACTA TAAAATGCA 60

TCGAATAATA CAA(GGCACTA GGCATGGTC TAGCAAAGGA AACAAACA ATG GGA GGT 116
Met Gly Gly
1

```

AGA GGT CGT GAG GCC AAA GTG GAA GTT CAA GGG AAG AAG CCT CTC TCA      164
Arg Gly Arg Val Ala Lys Val Glu Val Gln Gly Lys Lys Pro Leu Ser
         9          18          15

```

AGG GTT CGA AAC ACA AAG CCA TCA TTC ACT GTT GGC CAA CTC AAG AAA 212
 Arg Val Pro Asn Thr Lys Pro Pro Phe Thr Val Gly Gln Leu Lys Lys
 20 30 30 35

GCA ATT CCA CGA CAC TGC TTT CAG CGC TCC CTC CTC ACT TCA TTC TCC
Ala Ile Pro Pro His Cys Phe Gln Arg Ser Leu Leu Thr Ser Phe Ser
40 45 50

TAT GTT GTT TAT GAC CTT TCA TTT GCC TTC ATT TTG TAC ATT GCC ACC
Tyr Val Val Tyr Asp Leu Ser Phe Leu Ala Phe Val Pro Tyr Ile Asp Thr
15 60 65

Thr Tyr Pro Asn Leu Pro Asp Tyr Tyr Ser Glu Ile Val

ATC TAT TGG GTT CTC CAA GGT TGG CTT CTC ACT GGT GTG TGG GTG APT
 Ile Tyr Trp Val Leu Glu Tyr Tyr Leu Leu Thr Glu Val Tyr Val Ile
 36 36 36

GCT CAC GAG TGT GGT CAC CAT GCC TTC AGC AAG TAC CAA TGG GTT GAT 457
 Ala His Glu Tyr Gly His Ala Ile Ser Lys Tyr Glu Thr Val Asp
 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

GAT GTT GTC GGT TTG ACC CTT CAC TCA ACA CTT TTA GTC CCT TAT TTC		500
Asp Val Val Gly Leu Thr Leu His Ser Thr Leu Leu Val Pro Tyr Phe		
120	125	130
TCA TGG AAA ATA AGC CAT CGC CGT CAT CAC TCC AAC AAG GGT TCC CTT		548
Ser Trp Lys Ile Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu		
135	140	145
GAC CGT GAT GAA GTG TTT GTC CCA AAA CCA AAA TCC AAA GTT GCA TGG		596
Asp Arg Asp Glu Val Phe Val Pro Lys Pro Lys Ser Lys Val Ala Trp		
150	155	160
TTT TCC AAG TAC TTA AAC AAC CCT CTA GGA AGG GCT GTT TCT CTT CTC		644
Phe Ser Lys Tyr Leu Asn Pro Leu Gly Arg Ala Val Ser Leu Leu		
165	170	175
GTC ACA CTC ACA ATA GGG TGG CCT ATG TAT TTA GCC TTC AAT GTC TCT		692
Val Thr Leu Thr Ile Gly Trp Pro Met Tyr Leu Ala Phe Asn Val Ser		
180	185	190
195		
GGT AGA CCC TAT GAT AGT TTT GCA AGC CAC TAC CAC CCT TAT GCT CCC		740
Gly Arg Pro Tyr Asp Ser Phe Ala Ser His Tyr His Pro Tyr Ala Pro		
200	205	210
ATA TAT TCT AAC CCT GAG AGC GPP CTC ATC TAT GTC TCT GAT GTT GCT		788
Ile Tyr Ser Asn Arg Gln Asn Leu Leu Ile Tyr Val Ser Asp Val Ala		
215	220	225
TTG TTT TCT GTG ACT TAC CCT TAC CGT GTT GCA ACC CTG AAA GGG		836
Leu Phe Ser Val Thr Ser Leu Tyr Arg Val Ala Thr Leu Lys Gly		
230	235	240
TTG GTT TGG CTG CTA TGT GTT TAT GGG GTG CCT TTG CTC ATT GTG AAC		884
Leu Val Trp Leu Leu Cys Val Tyr Gly Val Pro Leu Leu Ile Val Asn		
245	250	255
GGT TTT CCT GTG ACT ATC ACA TAT TTG CAG CAC ACA CAC ITT GGC TTG		932
Gly Phe Leu Val Thr Ile Thr Tyr Leu Gln His Thr His Phe Ala Leu		
260	265	270
275		
CCT CAT TAC GAT TCA TCA GAA TGG GAC TGG CCT AAG GAA GCT TTG GCA		980
Pro His Tyr Asp Ser Ser Gln Trp Asp Trp Leu Lys Gln Ala Leu Ara		
280	285	290
VAL ATG GAC VAL GAT TAC TGT GTC TCA GAA ATG ATG TGT TAT TAT AIA		1026
Thr Met Asp Arg Asp Tyr Gly Ile Leu Asn Ile Val Phe His His Ile		
295	300	305
ACT GAT ATG GAT GAC GAC GAT CCT TGT TCA GAA ATG ATG TGT TAT AIA		
Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr		
310	315	320
CAT GCA AGC GAC GCA ACT ATT GCA ATC AAG CTA ATA TGT GGT GAC TAC		1124
His Ala Met Gln Ala Thr Asn Ala Ile Lys Pro Ile Leu Gly Gln Tyr		
325	330	335
TAC TAA TTT GAT GAC AGC CTA TTT TA' AAG TAA CTG TGT AGA GAA GCG		1177

Tyr Gln Phe Asp Asp Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala			
340	345	350	355
AGA GAG TGC CTC TAT GTG GAG CCA GAT GAA GGA ACA TCC GAG AAG GGC			1220
Arg Glu Cys Leu Tyr Val Glu Pro Asp Glu Gly Thr Ser Glu Lys Gly			
360	365	370	
GTG TAT TGG TAC ACG AAC AAG TAT TGATGGAGCA ACCAATGGGC CATACTGGGA			1274
Val Tyr Trp Tyr Arg Asn Lys Tyr			
375			
GTTATGGAAG TTGACTCATG TAATTAGTACA TAATTAGTAG AATGTTATAA ATAAGTGGAT			1334
TTGCCGCGTA ATGACTTTGT GTGTATTGTG AAACAGCTTG TTGCGATCAT GGTTATAATG			1394
AAAAAATAAT TCTCTTATTA ATTACATGTG CAAAGTGGTC TGCTTATAGC TTTCTGCCTA			1454
AAAAAAAAA			1462

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 379 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Gly Gly Arg Gly Arg Val Ala Lys Val Glu Val Gln Gly Lys Lys			
1	5	10	15

Pro Leu Ser Arg Val Pro Asn Thr Lys Pro Pro Phe Thr Val Gly Gln			
20	25	30	

Leu Lys Lys Ala Ile Pro Pro His Cys Phe Gln Arg Ser Leu Leu Thr			
35	40	45	

Ser Phe Ser Tyr Val Val Tyr Asp Leu Ser Phe Ala Phe Ile Phe Tyr			
50	55	60	

Ile Ala Thr Thr Tyr Phe His Leu Leu Pro Gln Pro Phe Ser Leu Ile			
65	70	75	80

Ala Tyr Pro Ile Tyr Tyr His Leu Leu Gln Lys Leu Leu Thr Gly Val			
85	90	95	100

Ile Leu Ile Asp His Gln Gly Asp His His Asp Phe Ser Lys Tyr Gln			
105	110	115	120

Trp Val Asp Asp Val Val Gly Leu Thr Leu His Ser Thr Ile Leu Val			
115	120	125	

Pro Tyr Phe Ser Trp Lys Ile Ser His Arg Arg His His Ser Asn Thr			
130	135	140	

Tyr Ser Leu Asp Asp Asp Gln Val Thr Val Pro Lys Pro Lys Ser Lys			
145	150	155	160

Val Ala Trp Phe Ser Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ala Val
 165 170 175
 Ser Leu Leu Val Thr Leu Thr Ile Gly Trp Pro Met Tyr Leu Ala Phe
 180 185 190
 Asn Val Ser Gly Arg Pro Tyr Asp Ser Phe Ala Ser His Tyr His Pro
 195 200 205
 Tyr Ala Pro Ile Tyr Ser Asn Arg Glu Arg Leu Leu Ile Tyr Val Ser
 210 215 220
 Asp Val Ala Leu Phe Ser Val Thr Tyr Ser Leu Tyr Arg Val Ala Thr
 225 230 235 240
 Leu Lys Gly Leu Val Trp Leu Leu Cys Val Tyr Gly Val Pro Leu Leu
 245 250 255
 Ile Val Asn Gly Phe Leu Val Thr Ile Thr Tyr Leu Gln His Thr His
 260 265 270
 Phe Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Lys Gly
 275 280 285
 Ala Leu Ala Thr Met Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe
 290 295 300
 His His Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met
 305 310 315 320
 Pro His Tyr His Ala Met Glu Ala Thr Asn Ala Ile Lys Pro Ile Leu
 325 330 335
 Gly Glu Tyr Tyr Gin Phe Asp Asp Thr Pro Phe Tyr Lys Ala Leu Trp
 340 345 350
 Arg Glu Ala Arg Glu Cys Leu Tyr Val Glu Pro Asp Glu Gly Thr Ser
 355 360 365
 Glu Lys Gly Val Tyr Trp Tyr Arg Asn Lys Tyr
 370 375

(2) INFORMATION FOR SEQ ID NO:7:

(i) POLYPEPTIDE CHAIN TERMINATION:
 (ii) LENGTH: 1720 base pairs
 (iii) ORIGIN: predicted by bioinformatics
 (iv) STRANDEDNESS: double
 (v) TOPOLOGY: linear

(vi) MOLECULE TYPE: cDNA

(vii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(v) ORGANISM: Zea mays

(vii) IMMEDIATE SOURCE:

(B) CLONE: pFad2#1

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 165..1328

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGCCTCTCC CCTCCCTCCT CCCTGCAAAT CCTGCAGACA CCACCGCTCG TTTTCTCTC	60
CGGGACAGGA GAAAAGGGA GAGAGAGGTG AGGGGGGGTG TCCGGGGGAT CTGCTCTGCC	120
CCGACGGAGC TGTTCAGTACC TCCTCAGTCT CAGTCAGGAG CAAG ATG GGT GCC GGC Met Gly Ala Gly	176
1	
GGC AGG ATG ACC GAG AAG GAG CGG GAG AAG CAG GAG CAG CTC GCC CGA Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu Gln Leu Ala Arg	224
5 10 15 20	
GCT ACC GGT GGC GCG ATG CAG CGG TCG CGG GTG GAG AAG CCT CGG Ala Thr Gly Gly Ala Ala Met Gln Arg Ser Pro Val Glu Lys Pro Pro	272
. 5 30 35	
TTC ACT CTG GGT TAC AAG AAG CGC ATC CGG CCA CAA TGC TTC GAG Ile Thr Leu Glu Ile Ile Lys Ala Ile Ile Ile Ile Lys Ile Ile Gly	310
40 45 50	
CGG TCG GTG CTC AAG TCC TTC TCG TAC GTG GTC CAC GAC CTG GTG ATC Arg Ser Val Leu Lys Ser Phe Ser Tyr Val Val His Asp Leu Val Ile	368
55 60 65	
GCG GCG CTC CGC TAC TTC GCG CTG GCG ATC ATA CGG GCG CTC CCA Ala Ala Ala Leu Tyr Phe Ala Leu Ala Ile Ile Pro Ala Leu Pro	416
70 75 80	
AGC CGG CTC CGC TAC GCG GCG TGG CGG TCG TAC TGG ATC CGG CAG GGG Ser Pro Leu Arg Tyr Ala Ala Trp Pro Leu Tyr Trp Ile Ala Glu Gly	464
85 90 95 100	
TGC GTG TGC ATC CGC CTG TGC CTC ATC CGG TAC GAG TGC GGC AAC CAC Cys Val Cys Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His	512
105 110 115	
CGG TTC TGG TAC TAC TCC CGT CTC TAC TAC TAC TGG TGT CGG TGC GTC His Asn Asn Asn Tyr Asn Ile Ile Asn Asn Asn Asn Asn Asn Asn	560
120 125 130 135	
CAC TCG TCG CGC ATG GTG CGC TAC TTC CGG TGG AAG TAC AGC CAC CGG His Ser Ser Leu Met Val Pro Tyr Pro Ser Trp Lys Tyr Ser His Arg	608
135 140 145	
CGC CAC CAC TCC AAC ACG GGG TCC CTG GAG CGC GAC GAG GTG TTC GTG Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val	656
150 155 160	
TAC AAC AAG AA GAG GCG CTG CGG TGG TAC ACC CGG TAT GTG TAC AAC	704

Pro Lys Lys Lys Glu Ala Leu Pro Trp Tyr Thr Pro Tyr Val Tyr Asn			
161	170	175	180
AAC CGG GTC GGC CCG GTG GTG CAC ATC GTG GTG CAG CTC ACC CTC GGG Asn Pro Val Gly Arg Val Val His Ile Val Val Gln Leu Thr Leu Gly			752
185	190	195	
TGG CTG CTG TAC CTG GCG ACC AAC GCG TCG GGG CGG CGG TAC CGG CGC Trp Pro Leu Tyr Leu Ala Thr Asn Ala Ser Gly Arg Pro Tyr Pro Arg			800
200	205	210	
TTC GCC TGC CAC TTC GAC CCC TAC GGC CCC ATC TAC AAC GAC CGG GAG Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr Asn Asp Arg Glu			848
215	220	225	
CGC GGC CAG ATC TTC GTC TCG GAC GGC GGC GTC GTC GGC GTG GCG TTC Arg Ala Gln Ile Phe Val Ser Asp Ala Gly Val Val Ala Val Ala Phe			896
230	235	240	
GGG CTG TAC AAG CTG GCG GCG GCG TPC GGG GTC TGG TGG GTG GTG CGC Gly Leu Tyr Lys Leu Ala Ala Ala Phe Gly Val Trp Trp Val Val Arg			944
245	250	255	260
GTG TAC GGC GTG CGG CTG CTG ATT GTG AAC GCG TGG CTG GTG CTC ATC Val Tyr Ala Val Pro Leu Leu Ile Val Asn Ala Trp Leu Val Leu Ile			992
265	270	275	
ATC TAC GTC GAC GTC AAC AGC GTC ATT TCA TCC TTT TAA GAC TCC TAC Tyr Tyr Leu Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser			1040
280	285	290	
GAG TGG GAC TGG CTG CGC GGC GCG CTG GGC ACC ATC GAC CGC GAC TAC Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met Asp Arg Asp Tyr			1088
295	300	305	
GGC AIC CTC AAC CGC GTG TTC CAT AAC ATC ACG GAC ACG CAT GTC CGG Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp Thr His Val Ala			1136
310	315	320	
CAC CAC CTC TTC TCC ACC ATG CGA GAC TAC CAC CAC GGC ATG GAG GGC ACC His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr			1184
325	330	335	340
AAG GCG ATC AGG CGC ATC CTC CGC GAC TAC TAC CAC TTC GAC CGG ACC Gly Ala Ile Arg Pro Ile Leu Gly Asp Tyr Tyr His Phe Asp Pro Thr			1232
345	350	355	
CGT GCG CGC AAC GAG GAG Gly Leu His Leu Ala Ile Ile Arg Gln Ala Gly Glu Cys Ile Tyr Val			
360	365	370	
GAG CGC GAG GAC CGC AAC GAG GAG GAG TTC TGG TAC GAG AAC AAC TGC TAC CGC CGC			
375	380	385	
CGCTGGCAGA GCTGAGGAGCCTTACCGTAGG AATGGAGCA GAAACGAGGA GGAGGAGAGG			1395
GTACTGCC GAAAGTCTCG GTCAGCTAT CTAATGTTA GTCGTGAGTC TTTTAGAGGG			1415

GAAGAAGAGAT CATTGGGCA CAGAGACGAA GGCTTACTGC AGTGCATCG CTAGAGCTGC	1515
CATCAAGTAC AAGTAGGCAA ATTTCGTCAAC TTAGTGTGTC CCATGTTGTT TTTCTTAGTC	1575
GTCGGCTGCT GTAGGCTTTC CGGGCGCGGT CGTTGTGTC GPTGGCATCC GTGGCCATGC	1635
CTCTGCCTGTC GTGGCGCGGC TTGTCTGTGTC CGTCTGTCGT CGCGTTGGCG TCGTCTCTTC	1695
GTGCTCCCCG TTGTGTTGTTG TAAAACAAGA AGATGTTTTC TGGTGTCTTT GGCAGAATAA	1755
CAGATCGTCC GAAAGAAAAAA AAAAAAAAAA AAAAAA	1790

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 387 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2;

Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu
1 5 10 15

25 25 30

Glu Lys Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro
35 40 45

His Cys Phe Glu Arg Ser Val Leu Lys Ser Phe Ser Tyr Val Val His
52 55 60

Asp Leu Val Ile Ala Ala Ala Leu Leu Tyr Phe Ala Leu Ala Ile Ile
 65 70 75 80

Pro Ala Leu Pro Ser Pro Leu Arg Tyr Ala Ala Trp Pro Leu Tyr Trp
35 90 95

Ile Ala Gln Gly Cys Val Cys Thr Gly Val Trp Val Ile Ala His Glu
100 105 110

Sys. Opt. File Missed Pkt. Sec. Avg. Min. Max. Min. Max. Min. Max. Min. Max.

Arg Lys Val Leu Val Ser Ser Ile Met Val Pro Tyr Phe Ser Trp Lys
333 344 355 366 377 388 399 400 411 422 433 444 455 466 477 488 499

Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Ileu Glu Arg Arg
145 150 155 160

Glu Val Phe Val Pro Lys Lys Lys Glu Ala Ileu Pro Trp Tyr Thr Pro
165 170 175

For Mail Type Auto Reset Dr. Max New Auto Max. Mail His Line 744-741-1000
2000 1980 1980

Leu	Thr	Leu	Gly	Trp	Pro	Leu	Tyr	Leu	Ala	Thr	Asn	Ala	Ser	Gly	Arg
195						200					205				
Pro	Tyr	Pro	Arg	Phe	Ala	Cys	His	Phe	Asp	Pro	Tyr	Gly	Pro	Ile	Tyr
210					215					220					
Asn	Asp	Arg	Glu	Arg	Ala	Gln	Ile	Phe	Val	Ser	Asp	Ala	Gly	Val	Val
225						230				235			240		
Ala	Val	Ala	Phe	Gly	Leu	Tyr	Lys	Leu	Ala	Ala	Ala	Phe	Gly	Val	Trp
					245				250				255		
Trp	Val	Val	Arg	Val	Tyr	Ala	Val	Pro	Leu	Leu	Ile	Val	Asn	Ala	Trp
					260			265				270			
Leu	Val	Leu	Ile	Thr	Tyr	Leu	Gln	His	Thr	His	Pro	Ser	Leu	Pro	His
					275			280			285				
Tyr	Asp	Ser	Ser	Glu	Trp	Asp	Trp	Leu	Arg	Gly	Ala	Leu	Ala	Thr	Met
				290		295				300					
Asp	Arg	Asp	Tyr	Gly	Ile	Leu	Asn	Arg	Val	Phe	His	Asn	Ile	Thr	Asp
305					310				315			320			
Thr	His	Val	Ala	Ala	His	Leu	Phe	Ser	Thr	Met	Pro	His	Tyr	His	Ala
					325			330			335			340	
Met	Glu	Ala	Thr	Lys	Ala	Ile	Arg	Pro	Ile	Leu	Gly	Asp	Tyr	Tyr	His
				345			350			355			360		
Phe	Asp	Pro	Thr	Pro	Val	Ala	Lys	Ala	Thr	Trp	Arg	Glu	Ala	Gly	Glu
				365		370			375			380			385
Cys	Ile	Tyr	Val	Glu	Pro	Glu	Asp	Arg	Lys	Gly	Val	Phe	Trp	Tyr	Asn
				390					395						
Lys	Lys	Phe													
		395													

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 671 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) NUCLEIC ACID TYPE: cDNA

(iii) SYNTHETICAL: No

(iv) ANTI-SENSE: No

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Ricinus communis

(vii) IMMEDIATE SOURCE:

(B) CLONE: pRFZ-1C

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..473

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGG	GTG	ATG	GCG	CAT	GAT	GTG	GCG	CAC	CAT	GCG	TTC	AGT	GAC	TAT	CAA		48		
Trp	Val	Met	Ala	His	Asp	Cys	Gly	His	His	Ala	Phe	Ser	Asp	Tyr	Gln				
1											10				15				
TTG	CTT	GAT	GAT	GTA	GTT	GAT	CTT	ATG	CTA	CAC	TCC	TGT	CTC	CTT	GTC		96		
Leu	Leu	Asp	Asp	Val	Val	Gly	Leu	Ile	Leu	His	Ser	Cys	Leu	Leu	Val				
																25	30		
CCT	TAT	TTT	TCA	TGG	AAA	CAC	AGC	CAT	GCG	CGA	CAT	CAT	TCC	AAC	ACA		144		
Pro	Tyr	Phe	Ser	Trp	Lys	His	Ser	His	Arg	Arg	His	His	Ser	Asn	Thr				
																35	40	45	
GGG	TCC	CTG	GAA	GGG	GAT	GAA	GTG	TTT	GTT	GCG	AAG	AAG	AAA	TCT	AGT		192		
Gly	Ser	Leu	Glu	Arg	Asp	Gly	Val	Phe	Val	Pro	Lys	Lys	Lys	Ser	Ser				
																50	55	60	
ATC	CGT	TGG	TAT	TCC	AAA	TAC	CTG	AAC	AAC	CCP	CCA	GGT	CGT	ATC	ATG		240		
Ile	Arg	Trp	Tyr	Ser	Lys	Tyr	Ile	Asn	Asn	Pro	Pro	Gly	Arg	Ile	Met				
																65	70	75	80
ACA	ATT	GCC	GTC	ACA	CTT	TCA	CTT	GCG	TGG	CGT	CTG	TAC	CTA	GCA	TTC		288		
Thr	Ile	Ala	Val	Thr	Leu	Ser	Leu	Gly	Trp	Pro	Leu	Tyr	Leu	Ala	Phe				
																35	30	35	
AAT	GTT	TCA	GGG	AGG	CCA	TAT	GAT	GGG	TTC	GCG	TGC	CAC	TAT	GAC	CCA		336		
Asn	Val	Ser	Gly	Arg	Pro	Tyr	Asp	Arg	Phe	Ala	Cys	His	Tyr	Asp	Pro				
																105	105	110	
TAT	GGG	CGG	ATG	TAC	AAT	GAT	CGG	GAG	CGA	ATC	GAG	ATA	TTC	ATA	TCA		384		
Tyr	Gly	Pro	Ile	Tyr	Asn	Asp	Arg	Gly	Arg	Ile	Gly	Ile	Phe	Ile	Ser				
																115	120	125	
GAT	GCT	GTT	GTT	GCT	GTC	ACT	TTT	GGT	CTC	TAC	CAA	CTT	GCT	ATA			432		
Asp	Ala	Gly	Val	Leu	Ala	Val	Thr	Phe	Gly	Leu	Tyr	Gln	Leu	Ala	Ile				
																130	135	140	
GCG	AAG	GGG	CTT	GCT	TGG	GTC	TGT	GTA	TAT	GGA	GTC	CCA	TTC	TTC			480		
*16	Tyr	Gly	Ile	Ile	Tyr	Ile	Ile	Tyr	Val	Ile	Asp	Val	Ile	Tyr	Ile				
																145	150	155	160
CGG	CGG	AAI	TCA	TTC	CGG	GGT	GAT	ACA	TTT	GTC	CAG	CAT	ACT	GAC			528		
Val	Val	Asn	Ser	Phe	Leu	Gly	Ile	Val	Tyr	Val	Asp	Val	Leu	Asn					
																165	170	175	
CCT	CCA	TTG	CCA	CAT	TAT	GAT	TGG	TGG	GAG	TGG	CTA	AGA	GGA				576		
Pro	Ala	Leu	Pro	His	Tyr	Asp	Ser	Ser	Gly	Tyr	Asp	Trp	Leu	Arg	Gly				
																180	185	190	
ATG	TTA	TTA	TTT	TTG	ATC	AGC	TTG	TTC			624								
ATG	TTA	TTA	TTT	TTG	ATC	AGC	TTG	TTC											

Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe
195 200 205

CAT AAC ATA ACG GAC ACT CAA GTA GCT CAC CAG CTT TTC ACC ATG CCC C 673
His Asn Ile Thr Asp Thr Gln Val Ala His His Leu Phe Thr Met Pro
210 215 220

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 224 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Trp Val Met Ala His Asp Cys Gly His His Ala Phe Ser Asp Tyr Gln
1 5 10 15

Leu Leu Asp Asp Val Val Gly Leu Ile Leu His Ser Cys Leu Leu Val
20 25 30

Pro Tyr Phe Ser Trp Lys His Ser His Arg Arg His His Ser Asn Thr
35 40 45

Gly Ser Leu Glu Asp Asp Glu Val Phe Val Pro Lys Dye Lys Ser Ser
48 53 58 63

Ile Arg Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Ile Met
65 70 75 80

Thr Ile Ala Val Thr Leu Ser Leu Gly Trp Pro Leu Tyr Leu Ala Phe
85 90 95

Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro
105 110 115

Tyr Gly Pro Ile Tyr Asn Asp Arg Glu Arg Ile Glu Ile Phe Ile Ser
118 123 128

Asp Ala Gly Val Leu Ala Val Thr Phe Gly Leu Tyr Gln Leu Ala Ile
130 135 140

Ala Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu
145 150 155 160

Val Met Asp Ser Asp Ile Val Ile Ile Ile Ile Ile Ile Ile Ile Ile
165 170 175 180

Pro Ala Leu Pro His Ile Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly
185 190 195

Ala Leu Ala Thr Val Asp Arg Arg Tyr Gly Ile Leu Asn Lys Val Phe
195 200 205

His Ile Ile Thr Asp Thr Ala Val Ala His His Leu Phe Thr Met Pro
210 215 220

(2) INFORMATION FOR SEQ ID NO:11:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1369 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

LIVE ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Echinus communis*

(vii) IMMEDIATE SOURCE:

(B) CLONE: pRF197-2-42

(ix) FEATURE:

(A) NAME / KEY: CDE

(B) LOCATION: 184-1547

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Phe	Gln	Gly	Cys	Ile	Leu	Thr	Gly	Leu	Trp	Val	Ile	Gly	His	Glu	Cys
				100				105					110		
GGC	CAT	CAT	GCT	TTT	AGT	GAG	TAT	CAG	CTG	GCT	GAT	GAC	ATT	CTP	GGC
Gly	His	His	Ala	Phe	Ser	Glu	Tyr	Gin	Leu	Ala	Asp	Asp	Ile	Val	Gly
				115				119					125		
CTA	ATT	GTC	CAT	TCT	GCA	CTT	CTG	CTT	CCA	TAT	TTT	TCA	TGG	AAA	TAT
Ile	Ile	Val	His	Ser	Ala	Leu	Leu	Val	Pro	Tyr	Phe	Ser	Trp	Lys	Tyr
				131				135					140		
AGC	CAT	CGC	CCG	CAC	CAT	TCT	AAC	ATA	GGA	TCT	CTG	GAG	CGA	GAC	GAA
Ser	His	Arg	Arg	His	His	Ser	Asn	Ile	Gly	Ser	Leu	Glu	Arg	Asp	Glu
				145				150					155		
GTG	TTC	GTC	CCG	AAA	TCA	AAG	TCG	AAA	ATT	TCA	TGG	TAT	TCT	AAG	TAC
Val	Phe	Val	Pro	Lys	Ser	Lys	Ser	Lys	Ile	Ser	Trp	Tyr	Ser	Lys	Tyr
				160				165					170		175
TTA	AAC	AAC	CCG	CCA	GGT	CGA	GTT	TTG	ACA	CTT	GCT	GCC	ACG	CTC	CTC
Leu	Asn	Asn	Pro	Pro	Gly	Arg	Val	Leu	Thr	Leu	Ala	Ala	Thr	Leu	Leu
				180				185					190		
CTT	GGC	TGG	CCT	TTA	TAT	TTA	GCT	TTC	AAT	GTC	TCT	GGT	AGA	CCT	TAC
Ile	Gly	Trp	Pro	Leu	Tyr	Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr
				195				200					205		
GAT	GGC	TTT	GCT	TGG	CAT	TAT	GAT	GTC	TAT	GGC	CCA	ATA	TTT	TCT	GAA
Asp	Asn	Leu	Pro	Val	Asp	Tyr	Asn	Leu	Tyr	Leu	Ile	Leu	Leu	Ser	Gly
				210				215					220		
AGA	AAA	AGG	CTT	CAG	ATT	TAC	ATT	GCT	GAC	CTC	GGA	ATC	TTT	GCC	ACA
Arg	Gly	Arg	Leu	Gln	Ile	Tyr	Ile	Ala	Asp	Leu	Gly	Ile	Phe	Ala	Thr
				225				230					235		
ACG	TTT	GTG	CTT	TAT	CAG	GCT	ACA	ATG	GCA	AAA	GGG	TTG	GCT	TGG	GTA
Thr	Phe	Val	Leu	Tyr	Gln	Ala	Thr	Met	Ala	Lys	Gly	Leu	Ala	Trp	Val
				240				245					250		255
ATG	CGT	ATC	TAT	GGG	GTG	CCA	TTG	CTT	ATT	GTT	AAC	TGT	TTC	CTT	GTT
Met	Arg	Ile	Tyr	Gly	Val	Pro	Ile	Ile	Val	Asn	Cys	Phe	Ile	Val	
				260				265					270		
ATG	ATC	ACA	TAC	TTG	CAG	CAC	ACT	CAC	CCA	GCT	ATT	CCA	CGC	TAT	GGC
Met	Ile	Thr	Tyr	Leu	Gln	His	Thr	His	Pro	Ala	Ile	Pro	Arg	Tyr	Gly
				275				280					285		
TCA	TTC	GAA	TGC	CAT	TGG	TTG	AAA	GCA	ATG	CTG	ACT	CTG	ATG	AAC	
Ser	Leu	Gly	Val	Asp	Gly	Leu	Leu	Gly	ATG	CTG	ACT	CTG	ATG	AAC	
				290				295					300		
GAT	TAT	GGC	CTG	TTG	AAT	AAA	GCA	TTG	CAT	AAU	ATU	CTA	GAC	ATC	TCF
Asp	Tyr	Gly	Val	Leu	Asn	Lys	Val	Leu	Phe	His	Asn	Ile	Ala	Asp	Thr
				305				310					315		
GTA	GCT	CAT	CAT	CTC	TTT	GCT	ACA	CAT	TAC	CAT	GCA	ATG	GAG		
Val	Ala	His	His	Leu	Phe	Ala	Thr	Val	Pro	His	Tyr	His	Aia	Met	Glu
				320				325					330		

GCC ACT AAA GCA ATC AAG CCT ATA ATG GGT GAG TAT TAC CGG TAT GAT Ala Thr Lys Ala Ile Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr Asp 340 345 350	1236
GGT ACC CCA TTT TAC AAG GCA TTG TGG AGG GAG GCA AAG GAG TGC TTG Gly Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Lys Glu Cys Leu 355 360 365	1284
TTC GTC GAG CCA GAT GAA GGA GCT CCT ACA CAA GGC GTT TTC TGG TAC Phe Val Glu Pro Asp Glu Gly Ala Pro Thr Glu Gly Val Phe Trp Tyr 370 375 380	1312
CGG AAC AAG TAT TAAAAAAAGTG TCATGTAGCC TGCCG Arg Asn Lys Tyr 385	1369

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 387 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ

(xi SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Gly Gly Gly Gly Arg Met Ser Thr Val Ile Ile Ser Asn Asn Ser

Glu Lys Lys Gly Gly Ser Ser His Leu Glu Arg Ala Pro His Thr Lys
 25 26 27 28 29 30 31

Pro Pro Tyr Thr Leu Gly Asn Leu Lys Arg Ala Ile Pro Pro His Cys
 35 40 45

Phe Glu Arg Ser Phe Val Arg Ser Phe Ser Asn Phe Ala Tyr Asn Phe
 30 35 40 45 50 55 60

Cys Leu Ser Phe Leu Ser Tyr Ser Ile Ala Thr Asn Phe Phe Pro Tyr
 61 71 75 83

Ile Ser Ser Pro Leu Ser Tyr Val Ala Trp Leu Val Tyr Trp Leu Phe
45 50 55

Glx Gly Cys Ile Leu Thr Gly Leu Trp Val Ile Gly His Glu Cys Gly

Arg Ser Val Asp Asp Asp Ile Val Gly Leu
106 107 108 109 110 111 112

Ile Val His Ser Ala Ile Ile Val Pro Tyr Phe Ser Thr Tyr Lys Tyr Ser
130 135 140

His Arg Arg His His Ser Asn Ile Gly Ser Leu Glu Arg Asp Glu Val
145 150 155 160

Asn	Asn	Pro	Pro	Gly	Arg	Val	Leu	Thr	Leu	Ala	Ala	Thr	Leu	Leu	Leu
						180			185						190
Gly	Trp	Pro	Leu	Tyr	Leu	Ala	Phe	Asn	Val	Ser	Gly	Arg	Pro	Tyr	Asp
						195			200						205
Arg	Phe	Ala	Cys	His	Tyr	Asp	Pro	Tyr	Gly	Pro	Ile	Phe	Ser	Glu	Arg
						210			215						220
Glu	Arg	Leu	Gln	Ile	Tyr	Ile	Ala	Asp	Leu	Gly	Ile	Phe	Ala	Thr	Thr
						225			230						240
Phe	Val	Leu	Tyr	Gin	Ala	Thr	Met	Ala	Lys	Gly	Leu	Ala	Trp	Val	Met
						245			250						255
Arg	Ile	Tyr	Gly	Val	Pro	Leu	Leu	Ile	Val	Asn	Cys	Phe	Leu	Val	Met
						260			265						270
Ile	Thr	Tyr	Leu	Gln	His	Thr	His	Pro	Ala	Ile	Pro	Arg	Tyr	Gly	Ser
						275			280						285
Ser	Glu	Trp	Asp	Trp	Leu	Arg	Gly	Ala	Met	Val	Thr	Val	Asp	Arg	Asp
						290			295						300
Tyr	Gly	Val	Leu	Asn	Lys	Val	Phe	His	Asn	Ile	Ala	Asp	Thr	His	Val
						305			310						320
Asn	Asn	Leu	Ser	Ile	Asn	Val	Arg	Asp	Ile	Asn	Ala	Met	Glu	Asp	
						325			330						335
Thr	Lys	Ala	Ile	Lys	Pro	Ile	Met	Gly	Glu	Tyr	Tyr	Arg	Tyr	Asp	Gly
						340			345						350
Thr	Pro	Phe	Tyr	Lys	Ala	Leu	Trp	Arg	Glu	Ala	Lys	Glu	Cys	Leu	Phe
						355			360						365
Val	Glu	Pro	Asp	Glu	Gly	Ala	Pro	Thr	Gln	Gly	Val	Phe	Trp	Tyr	Arg
						370			375						380
Asn	Lys	Tyr													
		385													

(Z) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHAIN FEATURES:

(A) LENGTH: 23 base pairs

(B) STRANDNESS: double

(C) POLYMER: nucleic acid

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION: 1..23
(D) OTHER INFORMATION: /product=
"synthetic
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGGGTATGCC ATGANTGGG NCA

23

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /product=
"synthetic
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AAARTGRTGC CAGTCGAGTC TC

22

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2973 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) ORIGINAL SOURCE:
The gene was isolated from human brain.

(vi) CLONING VECTOR:
(B) CLONE: pAGP2-6

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 433..520

23

(ix) FEATURE:

(A) NAME/KEY: intron
 (B) LOCATION: 521..1654

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATTCGGTAAT	TCCATACATAT	TTTGTAGAGATT	AGTTTGAGTT	TCCATCCATA	CTTTACTAGT	60
GATTATAAAAT	TTAAAATACG	TACTTTTCCA	CTATAAAAGTG	AAACTAAAGTA	AATTAGAACG	120
PGATATTAAA	AAGTTAATGT	TCACTGTTAT	ATTTTTTICA	CAAAGTAAAAAA	ATGGGGTTATT	180
TGGGGTAAAT	AAAAATACCA	GATAATTGGA	ATGATTAAA	AAGGGTGGAA	TAAGAGAGGA	240
GGGGAAAGAA	AAGAAGGTGG	GGGCCCCAGTA	PGAAAGGGAA	AGGTGTGTC	AAATCATCTC	300
TCTCTCTCTC	TACCTTGAC	CCACGGGCG	TGTCCATTAA	AAGGCCCTGTC	TCTTGCCATT	360
CCGCATCTGA	CCACGAGAAG	AAAGAGCACA	CACTCACAAA	TTAAAAAAAG	AGAGAGAGAG	420
AGAGAGACAG	AGAGAGAGAG	AGATTCTGCG	GAGGGAGCTTC	TTCTCTGTC	GCTGTTCTAC	480
GTTATTAAAG	TTATCGGCC	TACGTCAGCT	CCATCTCCAG	GGTCTGTC	TCTCTCCAT	540
TTCTTCTCAT	TTTGTGATT	TGTTCTTATT	TCTTTCCAGT	AGCTCTGCT	CTGTTGAATT	600
CTCGCGCTCAC	GATACATCTC	CTTATFACTCC	TTACATTCAA	CCTTGTAGATCT	GGTCTCGATT	660
TTTTTTTT	TTTTTTTT	TTTTTTTG	AAAWAAWAAW	GGGGGGGGAA	GGGGGGGGAA	720
CATTAATAAT	GATGAAATTCT	CTCAATTCTA	CAATGATTAG	TTTCTCTCGT	CTACCAAACG	780
ATATGTTGCA	TTTTCACTTT	TCTTCTTTTT	TTCTAAGATG	ATTTGCTTTG	ACCAATTGTT	840
TTAGATCTTT	ATTTTATTT	ATTTTCTGGT	GGGTTGGTGG	AAATTGAAAAA	AAAAAAAAAAA	900
AAAAGCATAA	ATTGTTATTT	GTTAATGTAT	TCATTTTTG	GCTATTGTT	CTGGGGTAAAG	960
ATCTGCTTCT	ACTGTTGAAT	CTTTCTGGA	TTTTTTACTC	CTATTGGTT	TTTATAGTAA	1020
AAATAACATAA	AAAAAGGAAA	ACAAAAGTTT	TATAGATTCT	CTTAAACCCC	TGAGGATAAA	1080
AGTTGCAATC	AAAATAATTC	AGGATCAGAT	GCTTTTGTGAT	GGATTCAGAT	GGGGATTACAG	1140
PTGCATGCAA	AATTTCCTAG	ATCCGTCGTC	ACATTTTATT	TTCTGTTAA	ATATCTAAAT	1200
CTGATATNTG	ATCTGCAA	ATTCTTCTT	CTTATACAT	ATTTGAGGAA	CTTCTTCTT	1260
GGTCTGTTG	TCATGTTGCT	TTTCTGATGAT	ATTTGCTGAT	TCGATGCTG	ATTTGTTATT	1320
ACAAAGCAAAC	TGATGTTAAC	CAAAAGCAAG	AGATGTGACCC	TGCGTTTATTA	ACATCGTATT	1380
ACCTTACTCT	AGTCTCTTTC	CGAAACGCAAT	CGTTTGTGTA	TTCTCTACAT	TATGGGGGCTT	1440
CTCTACTCTT	TATTCCTTTT	GGTCCACGCA	TTTTCTATT	GTGGCAATCC	CTTTCACAAAC	1500
CTGATTTCCC	ACTTTGGATC	ATTTGCTGAA	AGACTCTCTT	GAATCGTTAC	CACTTGTTC	1560
TTCTGCTT	TTTCTTCTT	AAATTAAATG	ATAAAACAT	TGATGTTCT	TGAGTGTGCA	1620

GCTTGTGAT	TCTTTGCTT	TTGGTTTCT	GCAGAACAT	GGGTGCAGGT	GGAAGAATGC	1689
CGGTTCTTAC	TTCTTCAAAG	AAATCGAAA	CGGACACCCAC	AAAGCGTGTG	CGCTGCGAGA	1743
AACCGCCTT	CTCGGTGGGA	GATCTGAAGA	AAGCAATCCC	GGCGATTGT	TCACAAACGCT	1800
CAAPOCCCG	CTCTTTCGCC	TACCTTATCA	GTGACATCAT	TATAGCCTCA	TGCTTCTACT	1860
ACGPOGCCAC	CAATTACTTC	TCTCTCTCC	CTCAGGCTCT	CTCTTACTTG	GCTTGGCCAC	1920
TCTATTGGGC	CTGTCAAGGC	TGTGTCTAA	CTGGTATCTG	GGTCATAGCC	CACGAATGCG	1940
GTCAACCACGG	ATTCAGCGAC	TACCAATGGC	GGATGACAC	AGTTGGTCTT	ATCTTCGATT	2040
CSTTCCTCT	CGTCGCTTAC	TTCTCTGGG	AGTATAGTC	TGCGCGTCAC	CATTCGAACA	2170
CTGGATCCCT	CGAAAAGAGAT	GAAGTATTG	TGCCAAAGCA	GAAATCAGCA	ATCAAGTGGT	2180
ACGGGAAATA	CCTCAACAAAC	CCCTTGGAC	GCATCATGAT	GTTAACCGTC	CAGTTTGTCC	2230
TCGGGTGGCC	CTTGTACTTA	GCCTTAAACG	TCTCTGGCAG	ACCGTATGAC	GGGTTCGCTT	2240
GCCATTTCCTT	CCCCAACGCT	CCCATCTACA	ATGACCGAGA	ACGGCTCCAG	ATATAACCTCT	2340
CTGATGCGGG	TATTCGAGCC	GTCTGTTTTG	GTCTTACCG	TTACGCTGCT	GCACAAAGGGA	2400
TGTTGTGAT	GATCTGCGTC	TAAGGAGTAC	CGCTTCTGAT	ATGAAATGG	TTCTCTGCT	2460
TGATCACTTA	CTTGCGACGAC	ACPCATGCCP	CGPTGCCTCA	CTACGATTCA	TCAGAGTGGG	2470
ACTGGCTTAC	GGGAGCTTG	GCTACCGTAG	ACAGAGACTA	CGGAAACGTC	AAACAAGGTGT	2540
TCCACAAACAT	TACAGACACA	CACGTGGCPC	ATCACCTGTT	CTCGACAATG	CGCGATTATA	2640
ACGAAATGGA	AGCTACAAAG	GCCTAAAGG	CAATTCTGGG	AGACTATTAC	CAGTTGATG	2700
GAACACCGTG	GTATGTGGCG	ATGATAGGG	AGGCAAAGGA	GTGATCTAT	GTAGAACCGG	2760
ACAGGGAAGG	TCGAAAGAAA	GGTGTGACT	GGTACAAACAA	TAAGTATGAA	GGATGATGCT	2820
GAAAGAAATG	TCGACTTTC	TCTTGTCTGT	TTGTCTTTG	TTAAAGAAGC	TATGCTTGT	2880
TTTAATAATC	TTATTGTCGA	TTTTGTTGTG	TTATGACATT	TTGGCTGCTC	ATTATGTTAT	2940

10.000 - 10.000 - 10.000 - 10.000 - 10.000

(A) LENGTH: 23 base pair

(5) TYPE: multicellular

ANSWER $\frac{1}{2} \times 2\pi \times 10^8 \times 10^{-12} = 10^6 \text{ rad/s}$

-11. PROPOSED TYPE.

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..23
- (D) OTHER INFORMATION: 'product= "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGGCATGTNG ARAANAGRTG RTG

23

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME KEY: misc_feature
- (B) LOCATION: 1..23
- (D) OTHER INFORMATION: 'product= "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGGCATGTC TAAANAGRTG RTG

23

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

Ala Ile Pro Arg His Cys

(4) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Ala Ile Pro Lys His Cys
1 5

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 4
- (D) OTHER INFORMATION: Xaa = Pro or Lys

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Ala Ile Pro Xaa His Cys
1 5

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 3
- (D) OTHER INFORMATION: Xaa = Ile or Ile

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Trp Pro Xaa Tyr Trp
1 5

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Trp Pro Leu Tyr Trp
1 5

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 3
- (D) OTHER INFORMATION: Xaa = Leu or Ile

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Trp Pro Xaa Tyr Trp
1 5

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Ala His Glu Cys Gly His
1 5

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Gly His Asp Cys Gly His
1 5

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 1
- (D) OTHER INFORMATION: Xaa = Ala or Gly

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 3
(D) OTHER INFORMATION: Xaa = Asp or Glu

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Xaa His Xaa Cys Gly His
1 5

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(P) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Leu Leu Val Pro Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(P) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Ile Leu Val Pro Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(P) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) ENDPOINT:

(A) NAME/KEY: unsure

(B) LOCATION: 1

(D) OTHER INFORMATION: Xaa = Leu or Ile

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Xaa Leu Val Pro Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Trp Lys Tyr Ser His Arg
1 5

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Trp Arg Ile Ser His Arg
1 5

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURES:

- (A) NAME/KEY: insure
- (B) LOCATION: ..
- (C) OTHER INFORMATION: Xaa = Arg or Lys

(ix) FEATURES:

- (A) NAME/KEY: insure
- (B) LOCATION: ..
- (C) OTHER INFORMATION: Xaa = Ile or Tyr

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Trp Xaa Xaa Ser His Arg
1 5

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Ser His Arg Arg His His
1 5

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Ser His Arg Thr His His
1 5

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) FEATURE:

- (A) NAME/KEY: unsure
- (P) LOCATION: 4
- (D) OTHER INFORMATION: Kaa = Arg or Thr

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Ser His Arg Kaa His His
1 5

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Thr Thr Thr Lys Glu
1 5

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Val Thr Tyr Leu His
1 5

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: insure
- (B) LOCATION: 1
- (D) OTHER INFORMATION: Xaa = Ile or Val

(ix) FEATURE:

- (A) NAME/KEY: insure
- (B) LOCATION: 5
- (D) OTHER INFORMATION: Xaa = Gln or His

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Xaa Thr Tyr Ile, Val
1 5

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Leu Pro His Tyr
1

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Leu Pro Trp Tyr
1

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 4
- (D) OTHER INFORMATION: Xaa = His or Trp

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Leu Pro Xaa Tyr

1

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 6
- (D) OTHER INFORMATION: Xaa = Arg or Lys

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Trp Leu Xaa Arg Ala Leu

1

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Gly Ser Asp Glu Ile Val

1

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: unsur..
(B) LOCATION: 1
(D) OTHER INFORMATION: Xaa = Trp or Tyr

(ix) FEATURE:
(A) NAME/KEY: unsur..
(B) LOCATION: 3
(D) OTHER INFORMATION: Xaa = Arg or Lys

(ix) FEATURE:
(A) NAME/KEY: unsur..
(B) LOCATION: 5
(D) OTHER INFORMATION: Xaa = Ala or Gly

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Xaa Leu Xaa Gly Xaa Leu
1 4

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Thr Val Asp Arg Asp Tyr Gly
1 5

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Thr Leu Asp Arg Asp Tyr Gly
1 5

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: unsure

- (B) LOCATION: 2
(D) OTHER INFORMATION: Xaa = Leu or Val

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Thr Xaa Asp Arg Asp Tyr Gly
1

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Thr His Val Ala His His Leu Phe
1

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Thr His Val Thr His His Leu Phe
1

(2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) FEATURE:
(A) LENGTH: 4 amino acids
(B) LOCATION: 4
(C) AMINO ACIDS: His, Ile, Val, Val

The sequence consists of: His, Ile, Val, Val.

Thr His Val Ile Val Val Phe
1

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

His His Leu Phe Ser Thr Met Pro His Tyr
1 10

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

His His Leu Phe Pro Gln Ile Pro His Tyr
1 10

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: unsure
(B) LOCATION: 6
(D) OTHER INFORMATION: Xaa = Pro or Ser

(ix) FEATURE:

(A) NAME/KEY: unsure
(B) LOCATION: 6
(D) OTHER INFORMATION: Xaa = Gln or Thr

(ix) FEATURE:

(A) NAME/KEY: unsure
(B) LOCATION: 7
(D) OTHER INFORMATION: Xaa = Ile or Met

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

His His Leu Phe Pro Gln Ile Pro His Tyr
1 10

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /product=
"synthetic
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

ATAGCCCCCC AA 12

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

TGGTCTTTTG GT 12

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /product=
"synthetic
oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CGAGATATGCG AGC 15

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 10 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION: 1..10
 - (C) OTHER INFORMATION: /product= "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GTCGACGAGG

10

(2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION: 1..12
 - (C) OTHER INFORMATION: /product= "synthetic oligonucleotide"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

AGATTTGCTTA CG

12

(2) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Ala Ile Pro Pro His Cys Phe
1 5

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 3
- (D) OTHER INFORMATION: Xaa = Leu or Pro

(iv) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 4
- (D) OTHER INFORMATION: Xaa = Ile or Leu or Val

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Ala Thr Xaa Xaa Tyr Trp
1 5

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

His Glu Cys Gly His
1 5